A COMPUTER PROGRAM FOR "FILTERING, SMOOTHING, EXTRAPOLATION IN DOSE-RESPONSE EXPERIMENTS WITH APPLICATION TO DATA ON RESPIRATORY TUMOR OF RATS"

by

Sylvia Campodónico

and

Jingxian Chen

GWU/IRRA/Serial TR-89/6
June, 1989

The George Washington University
School of Engineering and Applied Science
Institute for Reliability and Risk Analysis

Research Sponsored by

Contract N00014-85-K-0202
Project NR 042-372
Office of Naval Research

and

Grant DAAL 03-87-K-0056
U.S. Army Research Office

A COMPUTER PROGRAM FOR "FILTERING, SMOOTHING, EXTRAPOLATION IN DOSE-RESPONSE EXPERIMENTS WITH APPLICATION TO DATA ON RESPIRATORY TUMOR OF RATS"

by

Sylvia Campodónico

and

Jingxian Chen

Abstract

We develop a computer program for Kalman filtering, smoothing, extrapolation in dose-response experiments, based on the method proposed by Meinhold and Singpurwalla in 1987, and we apply this program to some published data on doses of bischloromethyl ether administered to rats, discussed in Chen and Singpurwalla (1989).

Key Words: Kalman Filtering, Kalman Filter Smoothing, Dose-Response Experiments.

1. Introduction

In a biological or an engineering system, we are often required to infer the response at a very low dose (or stress) level, based on the data of the observed values of responses at higher dose (or stress) levels. Meinhold and Singpurwalla (1987) have proposed a method involving a use of the technology of Kalman filtering and Kalman filter smoothing, for inference and extrapolations in certain dose-response, damage-assessment and accelerated - life-testing studies. Singpurwalla and Chen (1989) have discussed issues pertaining to a practical implementation of this methodology. Aiming at an easy application of the methodology to real cases, we have developed a computer program, which is described herein.

2. Algorithm

2.1 *Model* (Meinhold and Singpurwalla (1987))

Assume that $E(Y(x)) = \exp(-\alpha(x)\, x^{\beta(x)})$, where $\alpha(x)$, $\beta(x) > 0$ and "$E(Y(x))$" denotes "the expectation of response Y at dose (or stress) level x". Suppose that $Y^*(x) \overset{\text{def}}{=} \log(-\log Y(x))$ and $Y^*(x) \sim \mathcal{N}(\mu(x), \sigma^2(x))$. Then, the *observation equation* of the Kalman filter is

$$Y_j^* = F_j\, \theta_j + \nu_j, \tag{2.1}$$

the system equation of the Kalman filter is

$$\theta_j = \theta_{j-1} + w_j, \tag{2.2}$$

where $Y_j^* = Y^*(x_j)$, $\theta_j = [\gamma, \beta]^T_{x_j}$, $[1 \ \log x_j] = F_j$, $\nu_j = \nu(x_j)$, $w_j = w(x_j)$, $\gamma_j = \log \alpha(x_j)$, $\nu_j \sim \mathcal{N}(0, V_j)$ and $w_j \sim \mathcal{N}(0, W_j)$; ν_j is assumed independent of w_j; x_{j-1} is the dose (stress) previous to dose x_j.

2.2 *Computational Method* (Singpurwalla and Chen (1989)

Assume that $\theta_{j-1} \sim \mathcal{N}(\hat{\theta}_{j-1}, \Sigma_{j-1})$, then, upon observing Y^*_j, $\theta_j \sim \mathcal{N}(\hat{\theta}_j, \Sigma_j)$, where

$$\begin{aligned}
\hat{\theta}_j &= \hat{\theta}_{j-1} + K_j(Y_j - F_j\hat{\theta}_{j-1}), \\
K_j &= R_jF_j{}^T(F_jR_jF_j{}^T + V_j)^{-1}, \\
R_j &= \Sigma_{j-1} + W_j, \\
\Sigma &= (I - K_jF_j)\, R_j,
\end{aligned} \tag{2.3}$$

where I is an identity matrix. The relationship in (2.3) is referred to as *the forward recurrence equations* of the Kalman filter. *The backward recurrence equations* are

$$\begin{cases}
\hat{\theta}_j(T) = \hat{\theta}_j + J_{j+1}\, [\hat{\theta}_{j+1}(T) - \hat{\theta}_j], \quad \text{and} \\
\Sigma_j(T) = \Sigma_j - J_{j+1}\, [\Sigma_{j+1}(T) - R_{j+1}]\, J^T_{j+1},
\end{cases} \tag{2.4}$$

where $J_j = \Sigma_{j-1} R_j^{-1}$, $\hat{\theta}_j(T)$ is an estimation of θ_j based on all data, and we assume that $\theta_j \sim \mathcal{N}(\hat{\theta}_j(T), \Sigma_j(T))$ during backward recurrence.

Once $\hat{\theta}_T(T)$ is obtained, inference about $Y(x_{j+1})$ (x_{T+1} means a dose (stress) lower than x_T) is

$$E(Y(x_{T+1})) \approx \exp(- \exp\,(F_{T+1}\, \hat{\theta}_T(T))), \tag{2.5}$$

since $Y^*(x_{T+1}) \sim \mathcal{N}(F_{T+1}\, \hat{\theta}_T(T), F_{T+1}\, \Sigma_T(T)\, F^T_{T+1})$.

Besides, to implement the Kalman filter, we are required to give some necessary starting values, say starting doses (stresses) x(0) and x(-1), assumed

responses y(x(0)) and y(x(-1)) for getting θ_{x_0}; assume variances of $y^*(x(i))$ = log (-log y(x(i))), i=0, -1, and their correlation for calculate Σ_0. We also propose

$$W(x_j) = .5 * (1.5)^j (x_{j-1} - x_j)^2 \cdot \Sigma(x_0). \tag{2.6}$$

$$V(x_j) = (1/\exp(\exp(F_j\hat{\theta}_{j-1})))^2(1 - Y \exp(\exp(F_j\hat{\theta}_{j-1})))^2. \tag{2.7}$$

3. Usage of the Computer Program

The program has been written in Mikrosoft QuickBASIC version 4.0, to be run on an IBM Personal computer or "compatible". A printer should be available.

The filename of the program is tr896.BAS.

This program's flow chart is drawn up below.

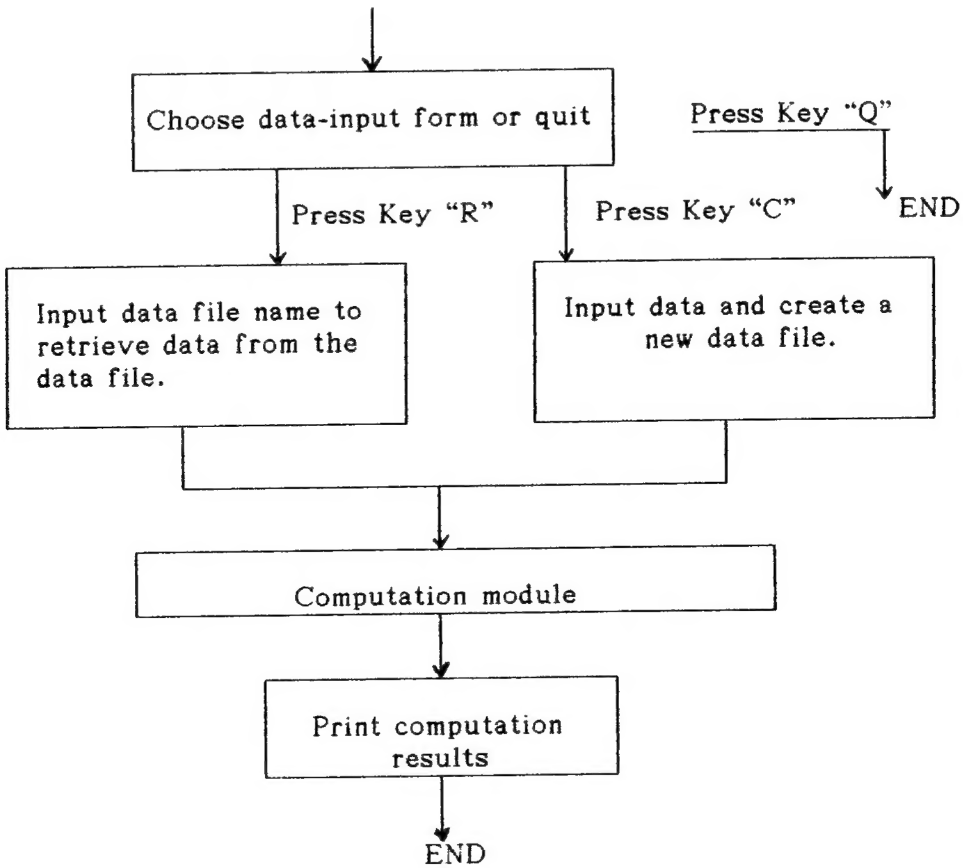

Figure 3.1 Flow Chart

This program contains a subroutine named "ENTRY" which allows either to input data and create a data file or to retrieve input from an existing file. As soon as the program gets started, "ENTRY" will be called. At this moment figure 3.2 appears on the screen.

```
  FILTERING, SMOOTHING, AND EXTRAPOLATIONS
       IN DOSE-RESPONSE EXPERIMENTS

An Input File is required for the calculations

PRESS  R  to retrieve input from existing file,

       C  to create new input file, or

       Q  to quit.
```

Figure 3.2 Selecting data-input form

There are three options:

1) to quit and return to the system, press key "Q";

2) to input data interactively and generate a data file, press key "C";

3) to retrieve input from an existing file, press key "R".

Once key "C" is pressed, the user should input the data requested on the screen (see figure 3.3). The input data will be saved in an ASCII data file.

Total number of pairs (stress level - response) considered by the experiment (ND) .... ?

Total number of pairs (stress level - response) for which predictions are required (NP) .... ?

Starting value of the stress level (x(0)) .... ?

	Stress level	Observed response
1. -	?	?
2. -	?	?
.	.	.
.	.	.
.	.	.
"ND." -	?	?

Stress levels for which responses are to be predicted

1. -	?
2. -	?
.	.
.	.
.	.
"ND." -	?

Second starting value of the stress level (x(-1)) .... ?
Assumed response at the starting stress level x(0) .... ?
Assumed response at the second starting stress level x(-1) .... ?

Let YO correspond to Y star of X subscript 0 on the paper

Assumed variance of YO .... ?

Let Y1 correspond to Y star of X subscript -1 on the paper

Assumed variance of Y1 .... ?
Assumed correlation of YO and Y1 .... ?
File name where input data will be saved .... ?

Figure 3.3 Hints for data input

Once key "R" is pressed, the user should enter the name of the data file to be retrieved, which format should meet the requirement of the program, i.e. the sequence of the data should be:

the number of data paris of observed responses and correspondent dose (stress) levels (ND),

the number of dose (stress) levels at which prediction is required (NP),

starting dose (stress) level (x(0)),

dose (stress) level 1(x(1)) and observed response at level x(1) (Y(x(11)),

dose (stress) level 2 (x(2)) and correspondent response (Y(x(2))),

. . . . . .

dose (stress) level ND (x(ND)) and correspondent response (Y(X(ND))),

X(ND + 1) - the 1st level at which prediction is required,

. . . . . .

X(ND+NP) - the last level at which prediction is required, another starting dose (stress) level (X(-1)) (note that X(i) < X(i-1), i=0, . . . , ND + NP, is required),

assumed responses at levels x(0) and x(-1) (Y(x(0)), Y(x(-1))),

assumed variances of ℓn(-ℓn Y(x(0))) and ℓn(-ℓnY(x(-1))),

assumed correlation of ℓn(-ℓnY(x(0))) and ℓn(-ℓnY(X(-1))).

After the user has either input all the data or retrieved the data from an existing file, the program prints all input data (as shown in Section 4, via an example) and starts computing. Finally, computation results are printed out. The print out is composed of three parts - estimation of parameters, dose-responses (including observed, predicted, filtered and smoothed responses) and prediction of

responses at low dose levels (as shown in Section 4, via the example.

4. An Application of the Program to Data on Respiratory Tumor of Rats

4.1 *A Group of Data on Respiratory Tumor of Rats and a Set-up of Starting Values*

The data on respiratory tumor of rats are abstracted from the Final Report of the SCFSC (1980), which are presented in Table 3.1.

Table 3.1

Dose-Response Data on Respiratory Tumor of Rats

Dose: x_i	10	20	40	60	80	100
Observed Response	40/41	43/46	14/18	14/18	19/34	8/20

In the table, datum "40/41" denote the fact that 40 out of 41 rats have not developed tumors.

Our starting values are given as following:

$x(0) = 250, \quad x(-1) = 280, \quad y(x(0)) = .005, \; y(x(-1)) = .001,$

$\mathrm{Var}\,(y^*(x(0))) = .001^2, \quad \mathrm{Var}\,(y^*(x(-1))) = .003^2, \; \rho(y^*(x(0)), y^*(x(-1))) = .7 .$

The doses at which response prediction is wanted, are 7, 5, 2, 1.

4.2 *The Printout of Input Data and Computation Results*

FILTERING, SMOOTHING, AND EXTRAPOLATIONS
IN DOSE-RESPONSE EXPERIMENTS

INPUT DATA SAVED AT "FILENAME$":

Total number of pairs (stress level - response) considered by the experiment (ND)	6
Total number of pairs (stress level - response) for which predictions are required (NP)	4
Starting value of the stress level (X(0))	250
Second starting value of the stress level (X(-1))	280
Assumed response at the starting stress level X(0)	1.667389
Assumed response at the second stress level X(-1)	1.932645
Let Y0 correspond to Y star of X subscript 0 on the paper. Assumed variance of ln(-ln(Y0))	.000001
Let Y1 correspond to Y star of X subscript -1 on the paper. Assumed variance of ln(-ln(Y1))	9E-08
Assumed correlation of ln(-ln(Y0)) and ln(-ln(Y1))	2.1E-07

Stress levels	Observed responses
100	.4
80	.5588
60	.7778
40	.7778
20	.9348
10	.9756

Stress levels for which responses are to be predicted:

7
5
2
1

From the starting conditions the following are the parameters

ALPHA, GAMMA, BETA and the covariance matrix of GAMMA and BETA, all of them at the starting stress level X(0):

START.ALF= 1.292872E-05	START.GAM=-11.25606	START.BET= 2.340585
SIGMA0: .0026858	8.486889E-05	-4.774242E-04

FILTERING, SMOOTHING, AND EXTRAPOLATIONS
IN DOSE-RESPONSE EXPERIMENTS
PROGRAM RESULTS

TABLE 1. Estimation of Parameters

DOSE	FILTERING		SMOOTHING	
	GAMMA	BETA	GAMMA	BETA
100	-11.24675	2.33893	-11.22904	2.335781
80	-11.21696	2.333635	-11.17025	2.325325
60	-11.14165	2.32025	-10.1507	2.144237
40	-9.543344	2.036153	-8.848135	1.912623
20	-8.113652	1.782019	-8.044925	1.769795
10	-7.605752	1.691735	-7.605752	1.691735

TABLE 2. Dose - Responses

DOSE	OBSERVED	PREDICTED	FILTERED	SMOOTHED
100	.4	.5376881	.5371238	.5360512
80	.5588	.6915619	.6898784	.6872347
60	.7778	.8271979	.8239535	.7758961
40	.7778	.9272031	.877179	.8466142
20	.9348	.9685547	.9395663	.9376571
10	.9756	.9820372	.9758289	.9758289

TABLE 3. Prediction of Responses at Low Levels

DOSE	PREDICTION
7	.986706
5	.992454
2	.998393
1	.999502

REFERENCES

Meinhold, R. J. and Singpurwalla, N. D. (1987). A Kalman-Filter Smoothing Approach for Extrapolations in Certain Dose-Response, Damage-Assessment, and Accelerated-Life-Testing Studies. *The American Statistician, Vol. 41, No. 2,* pp. 101-106.

Singpurwalla, N. D. and Chen, J. (1989). Filtering, Smoothing, and Extrapolations in Dose-Response Experiments with Application to Data on Respiratory Tumor of Rats. The George Washington University, Technical Report GWU/IRRA/Serial TR-89/5.

_____________ (1980). Final Report of the Scientific Committee of the Food Safety Counci, June 1980. Proposed System for Food Safety Assessment.